

SMOOTHNESS UNDER PARAMETER CHANGES: DERIVATIVES AND TOTAL VARIATION

Risto Holopainen

ABSTRACT

Apart from the sounds they make, synthesis models are distinguished by how the sound is controlled by synthesis parameters. Smoothness under parameter changes is often a desirable aspect of a synthesis model. The concept of smoothness can be made more accurate by regarding the synthesis model as a function that maps points in parameter space to points in a perceptual feature space. We introduce new conceptual tools for analyzing the smoothness related to the derivative and total variation of a function and apply them to FM synthesis and an ordinary differential equation. The proposed methods can be used to find well behaved regions in parameter space.

1. INTRODUCTION

Some synthesis parameters are like switches that can assume only a discrete set of values, other parameters are like knobs that can be seamlessly adjusted within some range. Only the latter kind of parameter will be discussed here. Usually, a small change in some parameter would be expected to yield a small change in the sound. As far as this is the case, the synthesis model may be said to have well behaved parameters.

A set of criteria for the evaluation of synthesis models were suggested by Jaffe [1]. Three of the criteria seem relevant in this context: 1) How intuitive are the parameters? 2) How perceptible are parameter changes? 3) How well behaved are the parameters? The vague notion of smoothness under parameter changes (which is not the name of one of Jaffe's criteria) can be made more precise by the approach taken in this paper.

From a user's perspective, the mapping from controllers to synthesis parameters is important [2]. In synthesis models with reasonably well behaved parameters, there are good prospects of designing mappings that turn the synthesis model and its user interface into a versatile instrument. However, a synthesis model does not necessarily have to have well behaved parameters to be musically useful. Despite the counter-intuitive parameter dependencies in complicated nonlinear feedback systems, some musicians are using them [3]. Likewise, acoustic instruments may have

far from smooth responses to changes in physical control variables (e.g. overblowing in wind instruments).

The smoothness of transitions has been proposed as a criterion for evaluating sound morphings [4]. As the morphing parameter is varied between its extremes, one would expect the perceived sound to pass through all intermediate stages as well. However, because of categorical perception some transitions may not be experienced as gradual. It may be impossible to create a convincing morph between, say, a banjo tone and a sustained trombone tone.

Quantitative descriptions of the smoothness of a synthesis parameter should use a measure of the amount of change in the sound, which can be regarded as a distance in a perceptual space. Similarity ratings of pairs of tones have been used in research on timbre perception, where multidimensional scaling is then used to find a small number of dimensions that account for the perceived distances between stimuli [5]. In several studies, two to four timbral dimensions have been found and related to various acoustic correlates, often including the attack time, spectral centroid, spectral flux and spectral irregularity [6]. The importance of spectrotemporal patterns was stressed in a more recent study [7] where five perceptual dimensions were found.

Most timbre studies have focused on pitched, harmonic sounds, in effect neglecting a large part of the possible range of sounds that can be synthesized. At the other extreme, the problem of similarity between pieces of music has been addressed in music information retrieval [8]. The difficulty in comparing two pieces of music is that they may differ in so many ways, including tempo, instrumentation, melodic features and so on. Most synthesis models of interest to musicians are also able to vary along several dimensions of sound, e.g., pitch, loudness, modulation rate and many timbral aspects. A thorough study of the perceived changes of sound would include listening tests for each synthesis model under investigation. A more tractable solution is to use signal descriptors as a proxy for such tests.

There are numerous signal descriptors to choose from [9], but the descriptors should respond to parameter changes in a given synthesis model. For example, in a study of the timbre perception of a physical model of the clarinet, the attack time, spectral centroid and the ratio of odd to even harmonics were found to be the salient parameters [10]. Since a synthesis model may be well behaved with respect to certain perceptual dimensions but not to others, the smoothness may be assessed individually for each of a set of complementary signal descriptors.

A synthesis model will be thought of as a function that

maps a set of parameter values to a one-sided sequence of real numbers, representing the audio samples. It will be assumed that all synthesis parameters are set at the beginning of a note event and remain fixed during the note. Dynamically varying parameters can be modelled by an LFO or envelope generator, but for simplicity we will consider only synthesis parameters that remain constant over time.

The effects of parameter changes may be studied either locally near a specific point in parameter space, or globally as a parameter varies throughout some range. The local perspective leads to a notion of the derivative of a synthesis model, which is developed in section 2. Parameter changes over a range of values are better described by the total variation, which is introduced in section 3. Then, sections 4 and 5 are devoted to case studies of the smoothness of FM synthesis and the Rössler attractor. Some applications and limitations of the methods are discussed in the conclusion.

2. SMOOTHNESS BY DERIVATIVE

In order to formalize the notion of smoothness, we will formulate a synthesis model explicitly as a function and describe what it means for that function to be smooth. First, we define a suitable version of the derivative. Then, in Sections 2.2 and 2.3, the practicalities of an implementation are discussed.

2.1 Definition of the derivative

Consider a synthesis model as a function $G : \mathbb{R}^p \rightarrow \mathbb{R}^N$ that maps parameters $\mathbf{c} \in \mathbb{R}^p$ to a one-sided sequence of samples x_n , $n = 0, 1, 2, \dots$, where the sample sequence will be notated $X(\mathbf{c})$ to indicate its dependence on the parameters. Then the question of smoothness under parameter changes is related to the degree of change in the sequence $X(\mathbf{c})$ as the point \mathbf{c} in parameter space varies. In practice, the distance in the output of the synthesis model will be measured through a signal descriptor rather than from the raw output signal. If a distance were to be calculated from the signals themselves, two periodic signals with identical amplitude and frequency but different phase might end up being widely separated according to the metric, despite sounding indistinguishable to the human ear. Signal descriptors that are clearly affected by the synthesis parameters and that can be interpreted in perceptual terms are preferable.

In order to treat the synthesis model as a function, it will be assumed to be deterministic in the sense that the same point in parameter space always yields identical sample sequences. The idea of relating how much a function $f(x)$ changes as the independent variable x changes by a small amount leads to the concept of derivative. Functions that have derivatives of all orders are called smooth. A more refined concept is to say that a function is k times continuously differentiable; the larger k is, the smoother the function.

Now, we would like to apply some suitably defined derivative to synthesis models considered as functions. To this end, a distance metric is needed for points in the parameter

space, and another distance metric is needed for points in the space of sample sequences. Let $d_p(\mathbf{c}, \mathbf{c}')$ be a metric in parameter space, and let $d_s(X(\mathbf{c}), X(\mathbf{c}'))$ be a metric in the sequence space. The derivative can then be defined as the limit

$$\lim_{\|\delta\| \rightarrow 0} \frac{d_s(X(\mathbf{c}), X(\mathbf{c} + \delta))}{d_p(\mathbf{c}, \mathbf{c} + \delta)} \quad (1)$$

where $\delta \in \mathbb{R}^p$ is some small displacement in parameter space. The limit, if it exists, is the derivative evaluated at the point \mathbf{c} .

In general, synthesis parameters do not make up a uniform space. Different parameters play different roles; they affect the sound subtly or dramatically and may interact so that the effect of one parameter depends on the settings of other parameters. This makes it hard to suggest a general distance metric that would be suitable for any synthesis model. Our solution will be to consider the effects of varying a single synthesis parameter c_j at a time, so the distance $d_p(\mathbf{c}, \mathbf{c}')$ in (1) reduces to $|c_j - c'_j|$. Furthermore, consider a scalar valued signal descriptor $\phi^{(i)}(\mathbf{c}) \equiv \phi^{(i)}(X(\mathbf{c}))$ which itself is a signal that depends on the sample sequence and the parameter value. Thus, we arrive at a kind of partial derivative evaluated with respect to the parameter c_j using a signal descriptor $\phi^{(i)}$,

$$\frac{\partial \phi^{(i)} \circ G(\mathbf{c})}{\partial c_j} = \lim_{h \rightarrow 0} \frac{d_s(\phi^{(i)}(\mathbf{c}), \phi^{(i)}(\mathbf{c} + h\mathbf{e}_j))}{h} \quad (2)$$

where \mathbf{e}_j is the j th unit vector in the parameter space. Clearly the magnitude of this derivative depends on the specifics of the signal descriptors used and which synthesis parameters are considered. In a finite dimensional space, all partial derivatives should exist and be continuous for the derivative to exist. Such a strict concept of derivative does not make sense in the present context where any number of different signal descriptors can be employed, so only the partial derivatives (2) will be considered.

Before discussing the implementation, let us recall some intuitive conceptions of the derivative. As William Thurston has pointed out [11], mathematicians understand the derivative in multiple ways, including the following.

- The derivative is the slope of a line tangent to the graph, if it has a tangent.
- In terms of symbolic operations, $\frac{d}{dx} x^n = nx^{n-1}$.
- The derivative is the best linear approximation to the function near a point.
- It is the limit of what you get by looking at a function under a microscope of higher and higher power.

Synthesis models are typically very complicated if considered as mathematical functions; hence the analytic approach to differentiation is out of the question and one has to rely upon numerical approximations. The various intuitions of what the derivative is may guide a practical numerical implementation in different directions, as will be further discussed in Section 2.3.

Numerical estimation of the derivative is highly sensitive to measurement noise. Here one source of measurement noise are the signal descriptors. Whereas one would like to magnify a curve in order to find its derivative at a point, doing so will also reveal more fine details caused by the noise, which may lead to false estimates. When properly estimated, the derivative will exaggerate irregularities and make them easier to detect.

2.2 Pointwise or time-average distance?

The distance metric d_s in sequence space has so far been left unspecified. We propose two alternatives, each suitable in different situations. The signal descriptors that will be used are based on short-time Fourier transforms of the signal $X(c)$ at regular intervals, using a hop size equal to the FFT window length, L . Hence, the signal descriptor is a sequence which we write concisely as $\phi_m(c)$, where $m = \lfloor n/L \rfloor$ is a time index.

Using a pointwise distance metric, one may follow the two signals over time and take the sum over their distances $|\phi_m(c) - \phi_m(c')|$ at each moment. Since these are infinite sequences, the sum may not converge. Therefore, an exponentially decaying weighting function is applied in the distance metric

$$d_s(X(c), X(c')) = \left[\sum_{m=0}^{\infty} \gamma^m (\phi_m(c) - \phi_m(c'))^2 \right]^{1/2} \quad (3)$$

where $\gamma \in (0, 1)$ controls the decay rate. Convergence is then guaranteed if the signal descriptors ϕ_m are bounded.

The second approach involves first taking an average over the sequence $\phi_m(c)$, $m = 0, 1, \dots, M$ and then comparing averages of two sequences. Thus, the distance becomes

$$d_s(X(c), X(c')) = |\langle \phi(c) \rangle - \langle \phi(c') \rangle| \quad (4)$$

where we take time averages

$$\langle \phi(c) \rangle = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=0}^{M-1} \phi_m(c) \quad (5)$$

before computing the distance. For time-varying signals, the drawback of the second approach is that two different temporal sequences ϕ_m may average to the same value.

As an illustration, consider two signals of equal average amplitude, the first having constant amplitude and the second with a periodic amplitude modulation. Suppose we compare the RMS amplitudes of the two signals using the second approach (4). When averaged over sufficiently long time, both signals will appear to have the same average amplitude. In contrast, the pointwise distance measure (3) will detect their difference.

2.3 Estimation of the derivative

A numerical computation of the derivative may return a number even if the limit (1) or (2) does not exist. Therefore, a measure of the reliability of the estimate, or “degree of differentiability”, should be added.

Although the synthesis model is assumed to be deterministic, all signal descriptors will introduce measurement noise. If a number of windowed segments of the signal are analyzed, then the spectrum of these segments will fluctuate unless some integer number of periods fit exactly into the window. The fluctuation can be reduced by using the time-averaged version of the distance metric (4).

Several methods for the estimation of derivatives exist [12]. Theoretically, it may be possible to arrive at analytical expressions for the derivative of a synthesis model considered as a function, at least in some trivial cases. In practice, numerical estimates have to be used. A simple approach would be to evaluate (2) directly at two points c and c' . Another approach is to fit a polynomial to the curve $\phi(c)$, and then do a symbolic differentiation of the polynomial.

The method of estimation of derivatives that will be used here is similar to one described in ref. [12, p. 231] but slightly simpler. The derivative at a point c_0 is approximated by a sequence of symmetric differences with decreasing distance h . A linear regression of this sequence gives the derivative as the intercept. Suppose a sequence of slopes

$$y_i(c_0; h_i) = \frac{\phi(c_0 + h_i) - \phi(c_0 - h_i)}{2h_i} \quad (6)$$

are given. Then the limit as $h \rightarrow 0$ can be found as the y-intercept of the fitted line

$$y_i = d + bh_i + \eta_i, \quad (7)$$

which gives the estimated derivative d . This method also provides a hint about the badness of fit, for which the root mean square error (RMSE) of the residuals η can be used.

3. TOTAL VARIATION

Whereas the derivative is concerned with local behaviour of a function, an even more useful perspective on the smoothness of a synthesis model may be to look at its properties over intervals of a parameter. One possible way to do so is to measure the length of the curve that a signal descriptor traces out as the parameter traverses some interval. If this curve is highly wrinkled, the curve becomes rather long, whereas a straight line connecting the endpoints means that the parameter changes are smooth. The total variation of a function may be used for such a measure; intuitively, it measures the length travelled back and forth on the y-axis of a function $y = f(x)$, $x \in [a, b]$.

Let $f(x)$ be a real function defined on an interval $x_0 \leq x \leq x_k$, and suppose $x_0 < x_1 < \dots < x_k$ is a partition of the interval. Then the total variation of $f(x)$, $x_0 \leq x \leq x_k$ is defined as

$$\mathcal{V}_{x_0}^{x_k}(f) = \sup \sum_{j=1}^k |f(x_j) - f(x_{j-1})| \quad (8)$$

taking the supremum over all partitions of the function. If f is differentiable, the total variation is bounded and can be expressed as

$$\mathcal{V}_{x_0}^{x_k}(f) = \int_{x_0}^{x_k} |f'(x)| dx. \quad (9)$$

Also, recall that one way for a function to fail to be differentiable is that its total variation diverges to infinity.

The mesh of the partition, which is the greatest distance $|x_j - x_{j-1}|$, needs to be fine enough when estimating the total variation numerically. A global description of the function's smoothness is obtained from considerations of the limit of the total variation as the mesh gets finer. Suppose the partition of $[x_0, x_k]$ is uniform with each point separated from its nearest neighbours by $|x_j - x_{j-1}| = \Delta$. Then, the question is whether a limit exists as $\Delta \rightarrow 0$.

For the present purposes it will suffice to consider approximations of the total variation using a small but fixed mesh. Certain functions may appear to have different amounts of total variation when observed at different scales. A slow increase in total variation as the mesh is successively made finer indicates that the estimation process goes as intended.

An alternative to measuring the total variation would be to measure the arc length, which can be thought of as the length of a string fitted to the curve if it is continuous. Fractal curves on the plane have the property that their arc length grows as the measurement scale gets smaller.

When measuring the total variation of a signal descriptor over a range of synthesis parameter values, there are still two possible approaches to how the distance is measured. As discussed above in section 2.2, either a pointwise distance may be taken, or the distance may be taken over time averages of the signal descriptors. The latter approach will be used here because it is better suited for the case of static parameters. Applications of the derivative and total variation to two synthesis models will be demonstrated next.

4. FM SYNTHESIS

With only three synthesis parameters, basic FM synthesis is convenient for investigations of the smoothness of its parameter space. The formula that will be used is

$$x_n = \sin(2\pi f_c n / f_s + I \sin(2\pi f_m n / f_s)) \quad (10)$$

with modulation index I , carrier frequency f_c , modulator frequency f_m and sample rate $f_s = 48$ kHz. Since the spectrum of the signal (10) is governed by a sum of Bessel functions [13], it may actually be possible to estimate some related signal descriptors directly from the formula, although we will not attempt to do so. The oscillations of the Bessel functions give FM synthesis its characteristic timbral flavour of partials that fade in and out as the modulation index I increases, with the overall brightness increasing with the modulation index. Brightness is related to the spectral centroid, which will be used to study the effects of parameter changes.

In the top of Figure 1, the centroid is shown as a function of I at two different carrier to modulator (C:M) ratios. The centroid, given in units of normalized frequency, is measured as the time average over 25 FFT windows using a

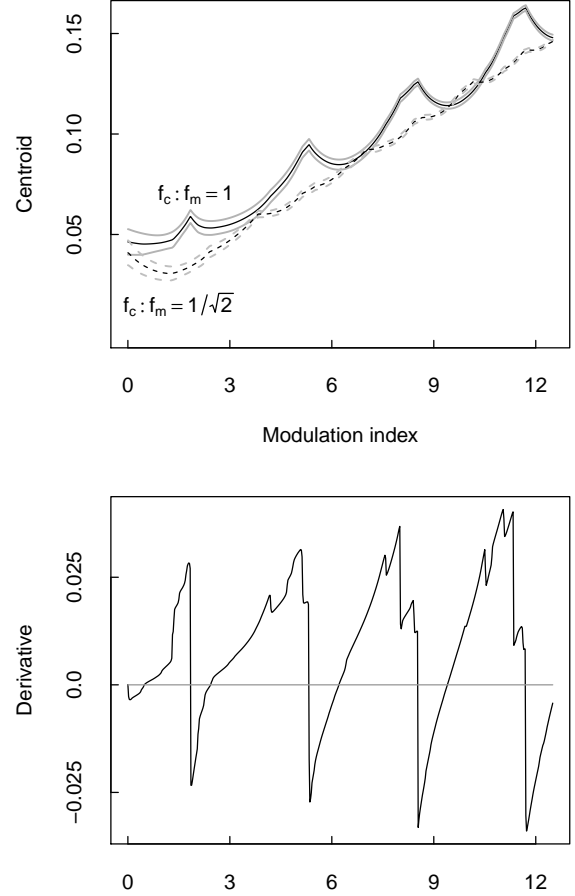


Figure 1. FM synthesis. Top: centroid as a function of modulation index for $f_c = f_m = 440$ Hz (solid line) and $f_c = 311.1$, $f_m = 440$ Hz (dashed line). The outer lines indicate one standard deviation of the centroid. Bottom: the derivative of the centroid at $f_c = f_m = 440$ Hz.

1024 point Hamming window. As can be seen, the C:M ratio 1 gives a rather bumpy curve with a general rising trend of the centroid, but with several local peaks. The bottom part shows the derivative, estimated with the method described in the end of Section 2.3. Evidently, the derivative is discontinuous at each of the peaks. The RMSE of the linear regression used in the estimation of the derivative is typically very small, but has sharp peaks around the discontinuities. It turned out to be necessary to re-initialize the oscillator's initial phase at the beginning of each run at a new parameter value, otherwise there would be oscillations in the centroid as a function of modulation index that would prevent the derivative from converging.

The total variation of the centroid over the range $0 < I \leq 12.5$ is about 0.127 for the inharmonic ratio $f_c/f_m = 1/\sqrt{2}$, and increases to about 0.188 for $f_c/f_m = 1$. We may now ask how the total variation changes as a function of the C:M ratio. This is shown in Figure 2. Narrow peaks arise at the simple C:M ratios 1 : 2, 1 and 3 : 2. Insofar as FM synthesis is reputed for its timbral variability as the modulation index varies, this phenomenon is more pro-

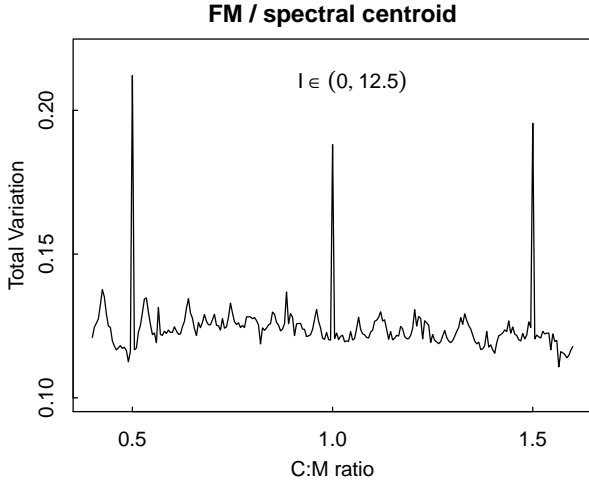


Figure 2. Total variation of the centroid of FM signals for $I \in [0, 12.5]$ as a function of the C:M ratio.

nounced at the simple C:M ratios that result in harmonic spectra.

Since the density of the spectrum depends on the modulation index as well as on the C:M ratio, signal descriptors related to spectral density may provide additional insights. The spectral entropy will be used for this purpose. Spectral entropy is measured from the amplitude spectrum, normalized so that all bins a_k sum to 1. Then, the normalized entropy is

$$H = -\frac{1}{\text{norm}} \sum_k a_k \log a_k \quad (11)$$

where a perfectly flat spectrum yields the maximum spectral entropy $H = 1$, and a sinusoid results in the smallest possible entropy of a signal that is not completely silent.

In Figure 3, the spectral entropy is shown as a function of the C:M ratio as well as the modulation index. Despite an even geometric progression of the modulation index $I \in [0.25, 20]$, the curves are slightly irregularly distributed. Two dips in spectral entropy can be seen at the simple ratios $C : M = 1, 2$. These dips can be understood to result from the fact that, at harmonic C:M ratios, several partials overlap (negative frequencies match positive frequencies), whereas for inharmonic ratios, there are more distinct partials in the spectrum.

The total variation of spectral entropy over the range of C:M ratios shown in Figure 3 is about 1 for $I = 0.25$, and it increases monotonically to a maximum value of 2.5 at $I = 1.25$. For higher modulation indices, the total variation decreases. These results can be interpreted as indicating that, if the modulation index is set at a fixed value and the C:M ratio is varied, then the sounds will change less for low modulation indices, and the maximum change occurs for $I = 1.25$.

5. THE RÖSSLER SYSTEM

Ordinary differential equations with bounded and oscillating solutions are good candidates for sound synthesis.

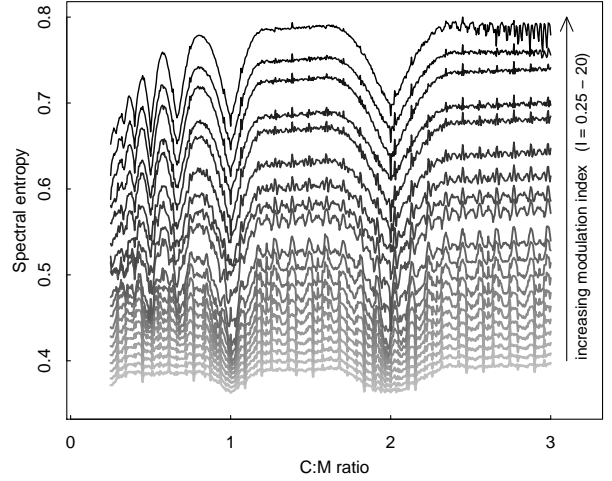


Figure 3. Spectral entropy of FM as a function of C:M ratio (horizontal) and modulation index (vertical).

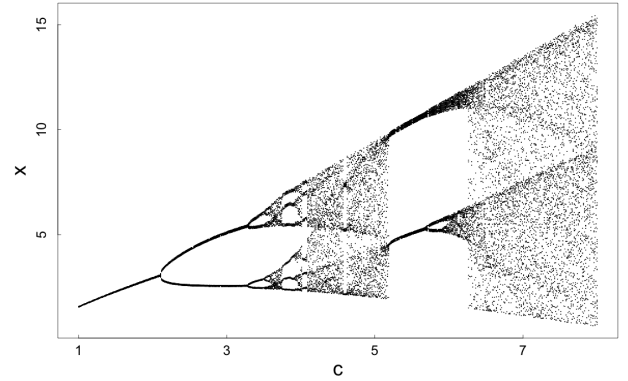


Figure 4. Poincaré section of the Rössler system showing bifurcations for $c \in [1, 8]$ and $a = b = 0.3$.

In particular, there are many nonlinear oscillators capable of both chaotic and periodic behaviour. Rössler's system [14],

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \quad (12)$$

is known to have a chaotic attractor at $a = b = 0.2, c = 5.7$. For lower values of c there are periodic solutions. A Poincaré section across the ray $x = -y, x \geq 0$ at $a = b = 0.3$ and a range of values of c reveals a period doubling route to chaos, after which there is a period two window (see Figure 4). In the following, (12) is solved with the fourth order Runge-Kutta method. The system is allowed time to approach an attractor by iterating at least 25000 time steps of size 0.025 before any measurements are taken.

The system rotates in the xy -plane, with occasional spikes in the z variable. Therefore, the x and y variables are suitable for use as audio signals, after they have been suitably

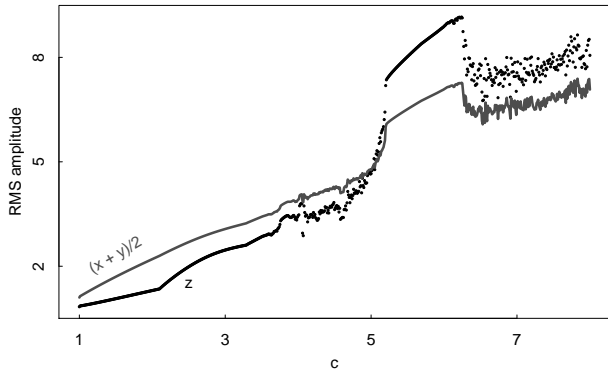


Figure 5. RMS amplitude of the Rössler system; the average of x and y is greater than z for low values of c .

scaled in amplitude. The first thing to check with an ordinary differential equation intended for use as an audio oscillator is its amplitude range and stability. As can be seen from Figure 4, the amplitude grows approximately linearly with c over the displayed range. By measuring the RMS amplitude of each coordinate, one gets a more detailed overview of the amplitude's dependence on the parameter c (see Figure 5). Because the amplitudes of x and y are typically not very different, their average has been plotted together with the amplitude of the z coordinate.

Bifurcation plots already reveal a few things about the smoothness under parameter changes. Each bifurcation is a point where the system's behaviour changes in a discontinuous way, whereas the behaviour between bifurcations can be expected to vary more smoothly.

Before going further, let us recall that dynamic systems may depend critically on the initial condition. Indeed, chaos is defined in terms of the exponential divergence of two orbits starting from infinitesimally separated initial conditions, which is measured with the largest Lyapunov exponent [15]. Even more dramatically, different initial conditions may lead to different kinds of behaviour. In conservative systems, orbits may be periodic, quasiperiodic or chaotic depending on the initial condition. Dissipative systems, such as Rössler's, have a basin of attraction of points that end up on the attractor, but should an orbit be started from outside the basin of attraction, it may wander off to infinity.

It is important to distinguish the properties of the orbit itself (chaotic versus regular) from the bifurcation scenarios as a parameter is varied. When looking at bifurcation diagrams, there are intervals of smooth change and intervals that are very irregular. It is tempting to guess that the irregular parts correspond to chaotic orbits, and the smooth parts to periodic orbits. This is only a half-truth; in fact, there are periodic windows interspersed with all the chaos.

As already seen, the RMS amplitude changes smoothly in some regions and irregularly in others. A quick comparison with the largest Lyapunov exponent λ indicates that the irregular parts correspond to chaotic regions (see Figure 6). Although it is easy to pick out "irregular regions" by visual inspection, a localized version of total variation can also achieve this. The local variation (LV) is defined as

the total variation over a short interval of length δ centred about a point x :

$$LV(f; x, \delta) = \mathcal{V}_{x-\delta/2}^{x+\delta/2}(f) \quad (13)$$

A mathematical definition of the LV would probably involve taking the limit $\delta \rightarrow 0$, but for practical purposes a small but finite interval must be used. Now the smoothness of a curve may be described in the neighbourhood of any point x_0 , which is computed by partitioning the interval into a suitably large number of points and proceeding as described above in Section 3. In the following example, $\delta = 0.02$ has been subdivided into 16 steps to find the local variation.

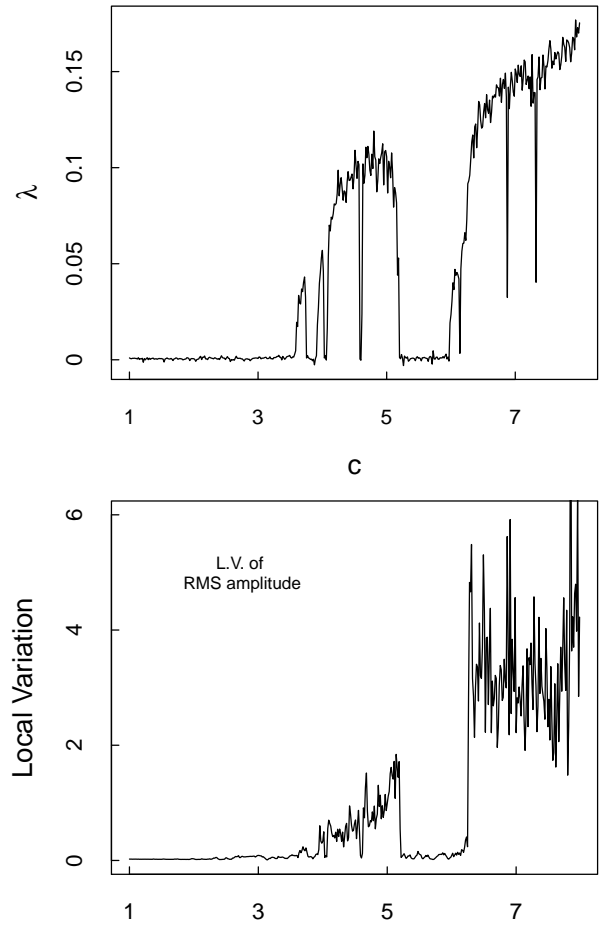


Figure 6. Greatest Lyapunov exponent (top) and local variation of the RMS amplitude (bottom) for the Rössler system as a function of the parameter c .

The local variation of the average RMS amplitude of the x and y coordinates of the Rössler system are shown in Figure 6 below a plot of the largest Lyapunov exponent over the same parameter range. When $\lambda = 0$, the dynamics is regular (either periodic or quasi-periodic), whereas $\lambda > 0$ indicates chaos. It is worth noting that regions of regular dynamics correspond to low values of the local variation, i.e., the amplitude changes smoothly. At chaotic regions, the local variation obtains higher values, although there is

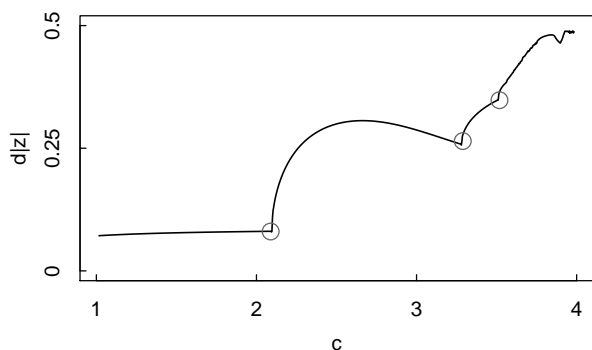


Figure 7. Derivative of the peak amplitude of the z coordinate as a function of c . Points of bifurcations are marked with circles.

no simple correlation between λ and LV. The higher values of LV in chaotic regions can be partly explained by the existence of periodic windows which may be very thin, yet are known to be dense in the chaotic regions.

In the interval $1 \leq c \leq 4$, there is a sequence of period doubling bifurcations. Most changes in amplitude are too subtle to notice directly (compare Figure 5), but taking the derivative, as shown in Figure 7, reveals points where the slope changes. In fact, the bifurcation points would be even easier to detect by plotting the second derivative of the peak amplitude.

In this study of the Rössler system, the effects of transients and dynamic parameter changes have been minimized. On the contrary, in a performance situation when using the Rössler system as an audio oscillator, its parameters would typically change over time. Then one may notice effects of hysteresis near bifurcations and in the chaotic regions. Approaching the same parameter value from different directions may then result in different behaviour.

6. CONCLUSION

By conceiving of a synthesis model as a function from points in parameter space to one-sided real sequences of audio samples, we have introduced a concept of derivative and total variation that can be used to describe the smoothness properties of the synthesis model. The derivative relates to local properties near specific points in parameter space, whereas the total variation characterizes the amount of change over intervals of a parameter. Interesting findings were that the total variation of the centroid with respect to the modulation index in FM synthesis is greater for simple harmonic C:M ratios than for other ratios. In other words, FM becomes smoother for inharmonic C:M ratios than for simple ratios. In the study of the Rössler system, we found that regular dynamics corresponds to smooth variation in the RMS amplitude. Chaotic regions are generally less smooth in parameter space, but there is some variation and relatively smooth parameter regions may exist where the system is chaotic as well.

The methods of characterizing the smoothness of synthesis models can be applied to analog synthesis and even to

acoustic instruments using mechanical transducers to excite them. Mechanical transducers may be needed also for the automated control of acoustic instruments by MIDI or other means, but the response characteristics of the transducer and the instrument considered together may not be known in advance and need to be mapped out. Analog, voltage controlled synthesizers can be similarly studied by applying some control voltage to one of its inputs. Then, studying the signal's response to changes in control voltage can further elucidate input to output relations and the smoothness of the parameter. Although smoothness properties can be roughly assessed by visual inspection, the derivative, and the total and local variations provide quantitative measures of smoothness.

Comparisons of smoothness properties across different synthesis models are, however, not so straightforward. One might intuitively want to argue that the Rössler system is less smooth, on the whole, than FM synthesis, but the set of synthesis parameters have entirely different meanings in the two models, so a direct comparison will be problematic. The same signal descriptors and distance metrics must of course be used for both synthesis models, and one must decide what parameter ranges to compare.

Noise is used in many kinds of synthesis. If the noise is prominent in the output signal, it will increase the variance of the signal descriptors and make the estimation of derivatives and total variation more complicated. If the noise is mild enough not to alter the behaviour of the synthesis model altogether, one can take ensemble averages over many runs of the system. Stochastic synthesis such as Xenakis' Gendyn algorithm [16] may however be beyond the scope of the present methods.

Ordinary differential equations and nonlinear feedback systems may exhibit hysteresis. In synthesis models with hysteresis, there is no longer a unique correspondence between the point in parameter space and the resulting output signal. This fact invalidates the assumption that the synthesis model can be thought of as a function that maps points in parameter space to sequences in the sample sequence space. Sometimes a transition from one type of behaviour to another may depend not only on the direction of the changing parameter, but also the speed of its change.

We began by making the assumption that signal descriptors could be used instead of conducting listening tests. This is obviously an exaggeration. Firstly, one needs to know what perceptual characteristics of sound are captured by various signal descriptors. Second, we have been looking at rather small variations in these descriptors and magnified them with the derivative or considered their total variation. It is very easy to gain a false impression that minor variations or roughnesses in the curves would be audible. Listening tests would be necessary in order to assess how the smoothness and irregularity of parameter changes are really perceived.

The assumption that maximally smooth parameters are always preferable is not necessarily true. Monotonicity and smoothness may be good, because then the parameter can be remapped in a way that is more practical for the user. Nevertheless, the rugged appearance of the param-

ter space of a chaotic system should not detract musicians from using them.

7. REFERENCES

- [1] D. Jaffe, “Ten criteria for evaluating synthesis techniques,” *Computer Music Journal*, vol. 19, no. 1, pp. 76–87, Spring 1995.
- [2] A. Hunt, M. Wanderley, and M. Paradis, “The importance of parameter mapping in electronic instrument design,” in *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, Dublin, Ireland, 2002.
- [3] D. Sanfilippo and A. Valle, “Towards a typology of feedback systems,” in *Proc. of the ICMC 2012*, Ljubljana, Slovenia, September 2012, pp. 30–37.
- [4] M. Caetano and N. Osaka, “A formal evaluation framework for sound morphing,” in *Proc. of the ICMC 2012*, Ljubljana, Slovenia, 2012, pp. 104–107.
- [5] J. Grey, “Multidimensional perceptual scaling of musical timbres,” *J. Acoust. Soc. Am*, vol. 61, no. 5, pp. 1270–1277, May 1977.
- [6] A. Caclin, S. McAdams, B. Smith, and S. Winsberg, “Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones,” *J. Acoust. Soc. Am*, vol. 118, no. 1, pp. 471–482, 2005.
- [7] T. Elliott, L. Hamilton, and F. Theunissen, “Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones,” *J. Acoust. Soc. Am*, vol. 133, no. 1, pp. 389–404, January 2013.
- [8] J.-J. Aucouturier and F. Pachet, “Music similarity measures: What’s the use?” in *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, Paris, France, October 2002.
- [9] G. Peeters, B. Giordano, P. Susini, N. Misdariis, and S. McAdams, “The timbre toolbox: Extracting audio descriptors from musical signals,” *J. Acoust. Soc. Am*, vol. 130, no. 5, pp. 2902–2916, November 2011.
- [10] M. Barthelet, P. Guillemin, R. Kronland-Martinet, and S. Ystad, “From clarinet control to timbre perception,” *Acta Acustica united with Acustica*, vol. 96, pp. 678–689, 2010.
- [11] W. Thurston, “On proof and progress in mathematics,” *Bulletin of the American Mathematical Society*, vol. 30, no. 2, pp. 161–177, April 1994.
- [12] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes. The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 2007.
- [13] J. Chowning, “The synthesis of complex audio spectra by means of frequency modulation,” *Journal of the Audio Engineering Society*, vol. 21, no. 7, pp. 526–534, September 1973.
- [14] O. Rössler, “An equation for continuous chaos,” *Physics Letters*, vol. 57A, no. 5, pp. 397–398, July 1976.
- [15] T. Tél and M. Gruiz, *Chaotic Dynamics. An Introduction Based on Classical Mechanics*. Cambridge University Press, 2006.
- [16] I. Xenakis, *Formalized Music. Thought and Mathematics in Music*. Stuyvesant: Pendragon Press, 1992.